10/588685
IAP20 Rec'd PCT/PTO 07 AUG 2006

Method for the calibration and verification of methylation analysis methods with the aid of non-methylated DNA.

Background of the invention.

The present invention relates to the use of DNA, in which 5-methylcytosine does not occur. Such non-methylated DNA is in particular required as a verification for a reliable and sensitive analysis of cytosine methylations.

5-methylcytosine is the most frequent covalently modified base in the DNA of eukaryotic cells. It plays an important biological role, amongst others for the transcription regulation, for the genetic imprinting and in the tumor genesis (for a survey: Millar et al.: Five not four: History and significance of the fifth base. In: The Epigenome, S. Beck and A. Olek (eds.), Wiley-VCH Verlag Weinheim 2003, pages 3 - 20). The identification of 5-methylcytosine as a component of genetic information is therefore of enormous interest. A detection of the methylation is however difficult, since cytosine and 5-methylcytosine have the same base pairing behavior. Many of the conventional detection methods based on hybridization can therefore not distinguish between cytosine and methylcytosine. Further, the methylation information gets completely lost with a PCR amplification.

The conventional methods for the methylation analysis are substantially based on two different principles. On the one hand, methylation-specific restriction enzymes are used; on the other hand, a selective chemical conversion of non-methylated cytosines into uracil (so-called bisulphite treatment, see for instance: DE 101 54 317 A1; DE 100 29 915 A1) takes place. The enzymatically or chemically pretreated DNA is then in most cases amplified and can be analyzed in different ways (for a survey: WO 02/072880 p. 1ff). Of great interest are methods, which are capable to sensitively and quantitatively detect methylation. Due to the important role of the cytosine methylation in the occurrence of cancer, this applies in particular with regard to diagnostic applications. Up to now, the conventional methods secure a sensitive and quantitative methylation analysis to a limited degree only.

For the sensitive analysis, the chemically pretreated DNA is usually amplified by means of a PCR method. By the use of methylation-specific primers or blockers, then a selective amplification only of the methylated (or in the reverse reaction: non-methylated) DNA is secured. The use of methylation-specific primers is known as the so-called "methylation-specific PCR" ("MSP"; Herman et al.: Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. Proc Natl Acad Sci USA. 1996 Sep 3; 93(18):9821-6). Thereby is achieved in particular the qualitative detection of methylated DNA having a low concentration. A comparably sensitive method is the so-called "HeavyMethyl" method. Therein, a specific amplification only of the originally

methylated (or non-methylated, resp.) DNA is achieved by the use of non-methylation-specific blocker oligomers (i.e. blocker oligomers, which hybridize at converted, originally non-methylated nucleic acids; for a survey: WO 02/072880; Cottrell et al.: A real-time PCR assay for DNA-methylation using methylation-specific blockers. Nucl. Acids. Res. 2004 32:e10). MSP as well as HeavyMethyl can be used as quantifiable "real time variants". They permit the detection of the methylation status of positions immediately in the course of the PCR, without a subsequent analysis of the products being required ("MethyLight" - WO 00/70090; US 6,331,393).

However, a reliable quantification of the methylation status over a linear range by the above methods is possible up to now to a limited degree only. For this, it is necessary that the assays are calibrated with fully methylated as well as with non-methylated DNA (cf.: Trinh et al.: DNA methylation analysis by MethyLight technology. Methods. 2001 Dec; 25(4): 456-62). The production of fully methylated DNA is relatively simple by the use of the SssI methylase. This enzyme transforms in the sequence context 5'-CG-3' all non-methylated cytosines into 5-methylcytosine. Problematic, however, is the production of fully non-methylated DNA. An enzyme corresponding to the SssI methylase, which quantitatively removes all methyl groups, is not available. Up to now, sperm DNA having a low methylation degree is used for the calibration (cf.: Trinh et al. 2001, ibid.). However, the sperm DNA is partially methylated and can thus to a limited degree only be used as a reliable standard. Further, artificially produced, short

non-methylated sequences such as PCR amplificates can also to a limited degree only be used, for instance for the analysis of individual defined positions. For multiplex reactions, these standards cannot be used, since the complexity of the reaction would then be too high. Further, the development of every new detection assay requires the production of a new defined standard. In contrast, a non-methylated standard covering the complete genomic DNA or a representative part thereof would permit a reliably quantifiable methylation analysis. Further, a standardized and thus simple, cost-effective and quick development of new detection assays would be possible. Because of the specific biological and medical importance of the cytosine methylation and because of the drawbacks mentioned above of the standards used today, there is a great technical need of methods, which make the genomic DNA in a fully non-methylated form available. In the following, such a - surprisingly simple - method is described.

According to the invention, so-called genome-wide amplification methods (WGA - whole genome amplification, for a survey: Hawkins et al.: Whole genome amplification - applications and advances. Curr Opin Biotechnol. 2002 Feb; 13(1): 65-7) are used for the production of non-methylated DNA. In this method, a large part of the genomic DNA is multiplied by means of a DNA polymerase and "random" or degenerated primers. "Random" primers are such primers, which do not specifically bind to certain nucleic acids, but to a multitude of nucleic acids. Thereto belong primers, which are either very short (between 5 and 10 bp), or primers, which are called "degen-

erated primers". Such degenerated primers are primers, which do not specifically bind to certain nucleic acids, since they contain either universal bases, which bind to several different nucleotides, or a mixture of primers is used, which differ in the "degenerated" positions. Universal bases are bases, which bind to several different nucleotides (see e.g. Promega catalog: Pyrimidine or purine-specific universal bases). Both are understood in the following as "degenerated primers". In the amplifications, only non-methylated cytosine triphosphates are offered, so that the amplificates are synthesized fully non-methylated. After several amplification runs, the amount of the partially methylated matrix DNA is fully in the background, compared to the newly produced, non-methylated nucleic acids.

Up to now, different WGA methods have been described. In the so-called primer extension preamplification (PEP), the amplification is performed by means of a Taq polymerase and a random mixture of oligonucleotide primers with a length of approx. 15 nucleotides (Zhang et al.: Whole genome amplification from a single cell: implications for genetic analysis. Proc Natl Acad Sci USA. 1992 Jul 1; 89(13): 5847-51}. In the DOP-PCR (degenerate oligonucleotide primed polymerase chain reaction), however, one degenerated primer only is used (cf.: Telenius et al.: Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. Genomics. 1992 Jul; 13(3): 718-25). Another WGA method is the so-called linker adaptor PCR. Therein, the DNA is first digested by means of a restriction enzyme. Then, linkers

are ligated at the restriction fragments. In the following amplification, primers are used, which specifically bind to the linkers (for a survey: Cheung and Nelson: Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA. Proc Natl Acad Sci USA. 1996 Dec 10; 93(25): 14676-9 with other documents). The PCR-based WGA methods described above have, however, several drawbacks. For instance, the generation of unspecific amplification artifacts may happen. Further, all genome sections are only incompletely covered. Moreover, there are created in part short DNA fragments having a length of less than 1 kB. (cf.: Dean et al.: Comprehensive human genome amplification using multiple displacement amplification. Proc Natl Acad Sci USA. 2002 Apr 16; 99(8): 5261-6 with other documents). The at present most effective method for the genome-wide amplification is thus the isothermal "multiple displacement amplification" (MDA, cf.: Dean et al. 2002 ibid.; US Patent 6,124,120). Therein, the genomic DNA is reacted with "random" primers and a DNA polymerase. Polymerases are used, which are capable to drive the non-template strand of the DNA double strand during the amplification out (for instance a φ29 polymerase). The driven-out strands in turn serve as a matrix for the extension of further primers. By this method it is possible to produce approx. 20-30 μg DNA from 1 - 10 copies only of human genomic DNA. This corresponds to a more than 5,000-fold amplification. The average product length is more than 10 kB, and the amplification occurs rather uniformly over the complete genome. The reaction may take place directly from biological

samples, for instance from blood or cell cultures. Commercially available kits for the MDA are at present offered by two suppliers ("GenomiPhi" of Amersham Biosciences, www.4amersham biosciences.com; "Repli-g" of Molecular Staging, www.molecularstaging.com). DNA already amplified is also available from these suppliers. The DNA produced by means of MDA is used in a great variety of applications, for instance in the genotyping of single nucleotide polymorphisms (SNP), in the "chromosome painting", in the restriction fragment length polymorphism analysis, in the subcloning and in the DNA sequentiation. The MDA can thus be used in particular for genetic, forensic and diagnostic investigations (cf.: Dean et al. 2002, ibid.).

The use of DNA produced by WGA methods as a standard in methods for the detection of 5-methylcytosine is not yet known up to now. The applications described in more detail in the following therefore allow the methylation analysis to have for the first time access to genomic, non-methylated DNA. Due to the special importance of the cytosine methylation and due to described drawbacks of the prior art, this advantageous, new technique is an essential technical progress.

Another aspect of the invention is the use of calibration standards containing non-methylated DNA for methods for the methylation analysis based on the use of microarrays, which are characterized by that detection oligomers are immobilized thereupon. According to the invention, therefore, a method based on the use of a microarray for the determination of the methyla-

tion degree of DNA under the use of calibration standards, which contain on the one hand non-methylated DNA and on the other hand specifically methylated DNA. "Specifically methylated" DNA is such a DNA, which is methylated to a known degree, i.e. has a known methylation rate. This method is characterized by that by means of the hybridization values, which are corrected for their noise, normalized and variance-stabilized in a multi-stage process, absolute values for the methylation rate can be obtained by means of a calculated calibration curve.

As a source of such DNA, sperm DNA may for instance be taken. Since the latter is however not fully non-methylated, it is preferred to use the already described genomic non-methylated DNA for this. Particularly preferred is the use of the genomic non-methylated DNA produced according to the described methods as a calibration standard. For this inventive aspect, it is however also imaginable to use non-methylated DNA from other sources for the calibration, as long as it is detectably non-methylated (e.g. sperm DNA).

Methods for the analysis of microarray experiments, wherein oligonucleotides for the detection of nucleic acids, such for instance mRNA or ESTs, or amplificates are immobilized on a surface, have been described. One problem of the analysis of gene expression microarray data is that the variation of the expression under constant conditions is not constant from gene to gene (David M. Rocke (2003) Heterogeneity of variance in Gene Expression Microarray Data. This publication is available as a preprint on

http://www.cipic.ucdavis.edu/-dmrocke/preprints.html, dated: March 15, 2003).

Another problem results however from the incomparability of expression microarray data from different microarray experiments, since they are very hard to calibrate. The approach to use so-called "housekeeping" genes as positive controls, permits a statement about whether a hybridization has in fact occurred (it is thus only tested whether the hybridization conditions were sufficient to permit hybridization of the analyte), but an absolute quantifiability of the expression data is not possible therewith. On the one hand, this is caused by that it is not clearly defined, which gene is under which conditions a "housekeeping gene", and on the other hand by that it is not possible to add a specific amount of known DNA, and to generate therewith an absolute value (for the calibration), since for this purpose it would be required to know before how many of the nucleic acids present in the sample to be investigated would actually bind to detection oligomer A, the respective amount of known DNA addition would then correspond to a signal intensity of 100% for exactly this detection oligomer. In other words, it is not clear, how much test DNA of a known sequence (standard) would have to be added to an experiment, in order to generate a value required for the calibration (i.e. corresponding to a specific, defined portion). Therefore, expression studies are always limited to relative statements. It cannot be determined, further, how large the portion of the mRNA expressed in the sample is, which hybridizes with a specific oligomer, e.g. oligo X, since the entirety of

signals does not need to correspond to the entirety of included mRNAs.

In the microarray methylation analysis mentioned here (see also Adorjan et al.: Tumor class prediction and discovery by microarray-based DNA methylation analysis. Nucleic Acids Res. 2002 Mar 1; 30(5): e21), oligomers for the detection of methylated (CG oligos) and/or non-methylated (TG oligos) CpG positions are immobilized. Frequently, dedicated pairs of CG and TG oligos are used and brought in relation in order to calculate a methylation index. Adorjan et al use log (CG/TG), Gitan et al. (Gitan et al. (2002) Methylation specific oligonucleotide microarray: a new potential for high throughput methylation analysis. Genome Res., 12, 158-164) use log(CG)/(log(CG)+log(TG). Thus estimations about, which CpG position of which sample is more or less methylated, are obtained. In order to find statistically significant markers, this method is sufficient. However, the real object is to determine true methylation rates, i.e. absolute values. This is in particular important for the following selection of the detection method, for the so-called assay format, in particular for the analysis of samples, which contain little DNA only. In the so-called "sensitive detection" methods being then applied, it is of importance, whether a low value corresponds to a 0-5% methylation rate or to a 15%-20% methylation rate.

For the mentioned microarrays it is known, which nucleic acids (amplificates) are included in the sample to be hybridized (not only whether they carry a C or a G in the sampling position),

and there are only two states, where they can be different, so that the entirety of the signals of a CG and a TG detection oligo is constant.

The method for the calibration of these microarray chips presented herein in an aspect of the invention has the advantage to permit for the first time access to absolute statements about the degree of methylation or the level of methylation of amplificates to be investigated.

It is particularly suited for the methylation analysis, which is characterized by that microarrays are used, at which at least two oligonucleotides per sampling position are immobilized in the amplificate to be detected. Further, it is characterized by that the nucleic acids (amplificates) to be hybridized against the oligonucleotides are per se known in their sequences, with the exception of one to three sampling positions, which in turn however can appear in two different variants only. These two detection oligos differ by that they have in the sampling position either a cytosine or a thymine, analogously, for the analysis of the counter strand, either a guanine or an adenine.

It is known that hybridization data of microarray hybridizations are first background-corrected and then normalized, further it is known that these data are converted, which is called data transformation, in order to obtain a variance stabilization. Variance-stabilized data are thus accessible to conventional statistical evaluations. The actual situation of the research in this field can for instance be taken

from the various publications of Durbin and Rocke or other articles mentioned therein:

- Rocke DM and Lorenzato S (1995), A two-component model for measurement error in analytical chemistry, Technometrics. 37, 176-184;

- Durbin BP et al. (2002), A variance stabilizing transformation for gene-expression microarray data. Bioinformatics. 18(1), 105-110;

- Geller SC et al. (2003), Transformation and normalization of oligonucleotide microarray data. Bioinformatics, 19(14), 1817-1823;

- Durbin BP and Rocke DR (2003), Estimation of transformation parameters for microarray data. Bioinformatics, 19(11), 1360-1367;

- Durbin BP and Rocke MR (2003), Approximate variance-stabilizing transformations for gene-expression microarray data, Bioinformatics, 19(8), 966-972;

- Durbin BP and Rocke MR (2004), Variance-stabilizing transformations for two-color microarrays, Bioinformatics, 20(5), 660-667.

Another aspect of the invention is however the step of the calibration of the microarray data under use of non-methylated DNA and thus also - now made accessible - DNA methylated to defined portions - i.e. specific -, as explained in more detail in the following.

Specification.

According to the invention, the DNA produced by genome-wide amplification methods is used as a standard in the methylation analysis. According to the invention is further provided a method for the methylation analysis, which is characterized by that

a) a genome-wide amplification is performed,

b) the amplificates received therefrom are used in the methylation analysis as a standard.

In principle, according to the invention, all WGA methods described above can be used. The reaction conditions of the PEP, DOP-PCR and linker-PCR also belong to the state of the art (see above). Because of the drawbacks of the PCR-based WGA methods, according to the invention an MDA is preferably performed. The reaction conditions for an MDA method are also sufficiently known (cf.: Dean et al 2002, ibid.; US patents 6,124,120; 6,280,949; 6,642,034; US application 20030143536; product information about the Genomiphi and Repli-g kits mentioned above). Other variations of the WGA, too, in particular of the MDA method, can be used according to the invention for the production of non-methylated DNA. For instance, it is possible to first fragment the DNA and to ligate linkers at the fragments. Subsequently, the fragments are transferred into concatamers, which are then amplified by means of an MDA (multiple strand displacement amplification of concatenated DNA MDA-CA; cf.: US 6,124,120).

According to the invention it is preferred to use a conventional MDA, however. Preferably, two

sets of primers are used. One primer set respectively is complementary to one strand of the DNA to be amplified. The primer sets may be random primers or degenerated primers. Details with regard to the number, length and structure of the primers have often been described (cf.: US 6,124,120). For instance, it is known that primers can be used, which are at the 5' end not complementary to the target sequence. Thereby, the driving-out of the primers by the polymerase is facilitated. The 5' region of the primers may in addition carry functional sequences, for instance for a promoter (cf.: US 6,124,120). The optimum structure of the primers depends on the type of the used polymerase, in particular on its processivity (cf.: US 6,124,120). Particularly preferred are hexamer primers. Different polymerases can be used in the MDA reaction. The enzymes must be capable either alone or in combination with auxiliary factors (for instance helicases) to drive the non-matrix strand of the DNA double helix to be replicated during replication out. For this, polymerases are preferred, which do not have a 5'-3' exonuclease activity. Alternatively, primers can also be used, which are blocked at the 5' end, and are therefore not degradable by the polymerases. As a polymerase, the φ29 polymerase is particularly preferred. The latter has a very high processivity permitting to synthesize DNA very effectively, even when extreme base compositions, short tandem repeats or secondary structures occur in the DNA. In the US patent 6,124,120 and in the US patent application 2003/0143536 A1, further possible polymerases are mentioned, such as Bst, Bca or phage M2-DNA polymerase. The reaction conditions required for the amplification depend on the se-

lection of the polymerases and the primers and also follow from the references named above. It is also known, amongst others, that a detection and quantification of the amplified DNA can be obtained by various methods, for instance by the incorporation of marked nucleotides, by the use of special detection probes or by solid phase detectors (cf.: 6,124,120).

In a preferred embodiment of the invention, the commercially available kits are used for the synthesis of the non-methylated DNA. Particularly preferred are the kits "GenomiPhi" (Amersham Biosciences) or "Repli-g" (Molecular Staging). The amplification takes place according to supplier's instructions. Basically, the DNA to be amplified is reacted with a sample buffer and random hexamer primers. The mixture is heat-denaturated and then cooled down, so that a binding of the primers to the DNA can occur. Then, the remaining reaction components, in particular the desoxynucleoside triphosphates and the φ29 polymerase are added. The reaction mixture is then incubated for approx. 30 hours at 30°C. As an initial material, for instance DNA can be used, which has been isolated by the commercially available purification methods. For cellular samples such as blood samples or primary cells from clinical samples, an alkaline lysis with subsequent neutralization may be sufficient (cf.: product information of Amersham for the GenomiPhi DNA amplification kit).

In a particularly preferred embodiment of the invention, commercially available DNA produced by means of MDA (see above) is used as a standard. This has the advantage that the DNA has a

constant concentration and quality because of the standardized production processes.

The DNA produced by using the above methods or commercially acquired can be used as a standard in a multitude of methylation analysis methods. Thereto belong methods based on the use of restriction enzymes as well as methods based on a bisulphite treatment of the DNA (cf.: Fraga and Esteller: DNA Methylation: A Profile of Methods and Applications. Biotechniques 33:632-649, September 2002). Preferably, first a bisulphite conversion is performed. The bisulphite conversion is known to the man skilled in the art in different variations (see for instance: Frommer et al.: A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci USA. 1992 Mar 1; 89(5): 1827-31; Olek, A modified and improved method for bisulphite based cytosine methylation analysis. Nucleic Acids Res. 1996 Dec 15; 24(24): 5064-6; DE 100 29 915; DE 100 29 915). It is particularly preferred that the bisulphite conversion is made in presence of denaturating solvents, such as dioxane, and a radical catcher (cf.: DE 100 29 915). In another preferred embodiment, the DNA is converted not chemically, but enzymatically. This is for instance imaginable by using cytidine deaminases, which react non-methylated cytidines more quickly than methylated cytidines. A respective enzyme has been identified just recently (Bransteitter et al.: Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. Proc Natl Acad Sci USA. 2003 Apr 1; 100(7):4102-7).

The converted DNA can be analyzed by means of conventional molecular biological methods, such as hybridization or sequentiation. In a preferred variant, the converted DNA is first amplified. For doing this, the man skilled in the art is familiar with different methods, such as ligase chain reactions. Preferably, the DNA is however amplified by a polymerase reaction. Various modifications are imaginable for this, for instance the use of isothermal amplification methods. Particularly preferred are however polymerase chain reactions (PCR). In a most particularly preferred embodiment, the PCR is performed by using primers, which specifically bind to positions of the converted sequence only, which were either methylated before (or in the reverse reaction: non-methylated) (MSP, see above). In another most particularly preferred embodiment, the converted DNA is analyzed by means of methylation or non-methylation-specific blockers ("HeavyMethyl" method, see above). The detection of the PCR amplificates may be made by conventional methods, for instance by methods of the length measurement such as gel electrophoresis, capillary gel electrophoresis and chromatography (e.g. HPLC). Mass spectrometry and methods for the sequentiation such as the Sanger method, the Maxam-Gilbert method and sequencing by hybridization (SBH) may also be used. In a preferred embodiment, the amplificates are detected by primer extension methods or by methylation-specific ligation methods (see for instance: Gonzalgo & Jones: Rapid quantitation of methylation differences at specific sites using methylation-sensitive single nucleotide primer extension (MsSNuPE). Nucleic Acids Res. 1997 Jun 15; 25(12): 2529-31; DE 100 10 282; DE 100 10

280). In another preferred embodiment, the amplificates are analyzed by means of hybridization at oligomer microarrays (cf.: Adorjan et al.: Tumor class prediction and discovery by microarray-based DNA methylation analysis. Nucleic Acids Res. 2002 Mar 1; 30(5): e21). In another particularly preferred embodiment, the amplificates are analyzed by using PCR real time variants (cf.: Heid et al.: Real time quantitative PCR. Genome Res. 1996 Oct; 6(10): 986-94, US patent No. 6,331,393 "MethyLight"). Therein, the amplification is performed in presence of a methylation-specific, fluorescence-marked reporter oligonucleotide. The reporter oligonucleotide then preferably binds to the DNA to be investigated and indicates the amplification thereof by increase or decrease of the fluorescence. It is particularly advantageous here to directly use the fluorescence change for the analysis and to infer a methylation state from the fluorescence signal. A particularly preferred variant is the "Taqman" method. In another particularly preferred embodiment, an additional fluorescence-marked oligomer is used, which hybridizes in immediate proximity to the first reporter oligonucleotide, and this hybridization can be detected by means of fluorescence resonance energy transfer ("Lightcycler" method).

It is a preferred embodiment of the invention to amplify several fragments at the same time by means of a multiplex PCR. Care has to be taken when designing that not only the primers, but also the other oligonucleotides used must not be complementary to each other, so that a high-degree multiplexing is more difficult in this case than usual. It is aggravating, further, that the

bisulphite-caused conversion of the nucleic acids reduces the complexity thereof. However, the chemically pretreated DNA offers the advantage that due to the differing G and C contents of the two DNA strands, a forward primer can never act as a reverse primer, too, which again facilitates the multiplexing and substantially compensates the disadvantage described above. The detection of the amplificates is in turn possible by different methods. The use of real time methods is for instance imaginable. For amplifications of more than four genes, it is however recommendable to detect the amplificates in a different way. An analysis by means of microarrays (see above) is preferred.

An updated survey of further possible methods for the methylation analysis is found in: Fraga and Esteller 2002, ibid.).

In the different methods for the methylation analysis, the MDA-DNA can be used as a standard in different ways. A standard is therein on the one hand any kind of negative control or positive control in the case of the detection of non-methylated DNA. This is in particular the case for technologies, which detect smallest amounts of methylated DNA in a big background of non-methylated DNA and vice versa. This case is also called "sensitive detection". Therein, the non-methylated MDA-DNA serves during the assay development as a verification of the specificity of the assay for methylated DNA and during the application of the assay as a negative control. It is however also preferred according to the invention to use a mixture of non-methylated DNA and methylated DNA. It is particularly preferred

to use different mixtures (i.e. consisting of different shares) of non-methylated and methylated DNA. Thereby, then calibration curves can be prepared. In order to prepare these mixtures, preferably the non-methylated DNA produced by MDA is used as a base. The total amount of the control DNA is subdivided and a part thereof is methylated by means of an SssI methylase (see above). The other part of the non-methylated DNA is also reacted with all reaction components of the methylation batch, except for the methylase. Thus it is secured that the DNA concentration in both batches is identical, and that in both batches the same reaction components are present. Subsequently, non-methylated and methylated DNA are mixed in different ratios, for instance in a ratio 4:0 for 0%, 3:1 for 25%, 2:2 for 50%, 1:3 for 75%, 0:4 for 100%. For the development of assays for the sensitive detection, it may be preferred to produce mixtures with very small concentrations of methylated DNA (for instance 1:2,000 - 1:10,000).

By calculating the quotient of the signals, which are detected for the methylated state, and the signals, which are detected for the non-methylated state, the measured methylation rate is obtained. If this is plotted against the theoretical methylation rates (according to the share of methylated DNA in the defined mixtures), and the regression through the measured points is determined, a calibration curve is obtained. By using this calibration curve, the methylation level of the unknown samples can be determined by means of the measured methylation rate.

For assays quantifying the methylation by hybridization of fluorescent measuring probe (microarrays, MethyLight, other hybridization assays in solution), the measured fluorescence signal is linearly correlated over a wide range with the concentration of methylated DNA (if the probe specifically hybridizes at - before the conversion - methylated nucleic acids) (JG Wetmur, Hybridization and renaturation kinetics of nucleic acids, Annual Reviews, 1976). These assays can therefore be calibrated by determining the unspecific background hybridization and the dynamic range of the measuring probes only. This may be achieved by means of artificially produced non-methylated and methylated DNA. The actual methylation rate can then be determined by simple linear interpolation between the fluorescence intensity of the non-methylated DNA (0% methylation) and the fluorescence intensity of the methylated DNA (100% methylation). It is particularly preferred that this determination of the methylation rate of a given sample is made for several repeated measurements and the statistical distribution of the measurement noise is taken into account (DM Rocke and S Lorenzato, A two-component model for measurement error in analytical chemistry, Technometrics, 1995, 37, 176-184). This is preferably implemented by means of classical statistical methods such as the "maximum likelihood" or "variance stabilization".

In the following, this particularly preferred method according to the invention for the conversion of signal intensities from methylation microarray hybridization experiments is described. Such methylation-specific microarray

hybridization experiments have already been described in detail elsewhere, for instance by Adorjan et al. (Nucleic Acids Res. 2002 Mar 1; 30(5): e21). The conversion of the signal intensities into methylation values is made by Adorjan et al. according to a different method, however. The new method is characterized by the use of non-methylated DNA according to the invention in an essential step, namely the calibration of the normalized and variance-stabilized hybridization data previously corrected for the background noise, said data occurring in particular in the methylation analysis with so-called CG and TG oligos. Only the calibration made possible by the use of non-methylated DNA permits at last the accurate assignment of methylation signals, which can be calculated from signal intensities, to actual methylation rates.

It is therefore another aspect according to the invention to provide a method, which is characterized by that it uses calibration standards for evaluating the hybridization data from microarray tests, wherein one standard does not contain methylated DNA, and the other standard contains DNA of a defined degree of methylation (e.g. 100%), in order to thus determine actual methylation rates. The methylated standard is preferably produced, as described above, by methylation with the SssI methylase. For use as a standard, which does not contain methylated DNA, preferably non-methylated DNA is used. Particularly preferred is herein the use of genomic non-methylated DNA produced according to the method described above.

Preferred, but not necessarily required for the preparation of a calibration curve is the use of several DNA's methylated in different degrees.

Particularly preferred is however a method, which permits to determine an accurate assignment of the absolute values by using a calibration curve, which has been calculated with the aid of two calibration standards only (particularly preferred are here 0% and 100%).

Therein, the methylation rate represents a value from 0 to 100, which indicates the ratio of the shares of the actually methylated DNA and of the non-methylated DNA with respect to each other in the analyzed sample per oligonucleotide probe pair (or per sampling position (= CpG or TpG), if the oligo covers one CpG site only).

According to the invention, the calibration standards are used for preparing a calibration curve, by means of which - as will be explained in more detail in the following - the actual methylation rates of the investigated samples can be read. According to the invention is provided, therefore, the specific method of the conversion of the obtained hybridization values into absolute methylation rates. This will be explained in more detail in the following.

Before the methylation data of a methylation experiment (based on microarrays) can be calibrated at all, the coarse PIXEL intensity statements of the microarray laser scan illustrations must be transferred for every oligonucleotide probe pair into methylation signal values. The

coarse measurement data of individual oligonucleotide spotting positions must therefore be converted into methylation signals per oligonucleotide probe pair (or per sampling position, resp.) of the detection probes. A sampling position is always the cytosine to be investigated (or thymine) before a guanine (within a CpG dinucleotide). Since detection oligonucleotide probes may however also contain several such CpG positions, methylation signals are indicated in the following per oligonucleotide probe pair and not per sampling position.

The evaluation method according to the invention is explained in more detail in the following:

The first step is the background noise correction: For every spot, i.e. every localized detection oligo, the median of the background pixel intensities is subtracted from the median of the foreground pixel intensities. This results in a robust estimation of the background-corrected oligonucleotide hybridization intensities. This may for instance also take place by using the formula: $X = Imeth- * r^2$.

The second step can be designated as normalization of data: For this purpose, generally classical methods are used, which can however be optimized for the application in methylation microarray experiments, by assuming that the sum of the CG signals and of the TG signals within a sample is constant, and that over several microarrays, or also over several amplificates only within a microarray. This is the inherent advantage of the methylation analysis with CG

and TG oligo probes, as already described elsewhere. The background-corrected redundant CG and TG detection oligo spot intensities of every microarray are linearly scaled, so that the overall distribution of the intensities of every microarray is as identical as possible (simplest case: median of the CG+TG intensities is identical for all microarrays). The intensities are preferably redundant, which means that spots (a defined applied quantity) of the same probe occur multiply on a microarray.

The third step is a transformation of data, which has the aim of variance stabilization. Variance stabilization is here the generation of variance-stabilized methylation signals in the generalized logarithmic space. Up to now, this has been achieved by that the logarithm of the ratio of measurement data generated by cytosine (CG) and thymine (TG) oligonucleotides (as described before by Adorjan et al. Nucleic Acids Res. 2002 Mar 1; 30(5): e21). The generated data should be variance-stabilized, in order that the conditions for the application of established statistical evaluation methods, such as the "maximum likelihood algorithm (ML)" are fulfilled. Instead of using the simple logarithm, here the generative model of Rocke is preferably used, which takes into account the specific noise (intensity-independent, caused by the use of fluorescence intensities), which is inherent in this measurement process. The use of a generalized log scale instead of the logarithm has the advantage that negative intensity values, as they may be generated by the background noise correction, can be taken into account, and very

low intensities are taken into account to a larger degree.

For a more detailed explanation of the generalized log scale, reference is made to the respective publications of Rocke in the magazine "Bioinformatics" (Durbin B and Rocke DR (2003) Estimation of transformation parameters for microarray data. Bioinformatics vol 19, no 11, pages 1360-1367).

Based on the use of negative control oligos, and the intensity data generated therewith, a global (i.e. valid at least for all oligonucleotide probe pairs immobilized on the microarray) background hybridization intensity distribution can be estimated. Negative control oligos are detection oligos, which are designed such that they do not hybridize at any of the amplificates to be investigated, which are known in these hybridization experiments for the methylation analysis in contrast to hybridization experiments for the expression analysis. The mean value (μ) and the variance ($\sigma 2$) for normal distribution (e.g. according to Gauss) are estimated.

A variance-stabilizing transformation first described by Durbin and Rocke (Durbin BP et al. (2002), A variance-stabilizing transformation for gene expression microarray data, Bioinformatics, 18(1), 105-110) is used to transform therewith CG and TG methylation measurement values (= intensities) per oligonucleotide probe pair into a methylation signal on a generalized log scale. As already mentioned, this generalized log scale permits a consistent treatment or

evaluation of low or even negative intensity signals. These methylation signals thus obtained fulfill therefore sufficiently the condition for the application of the conventional standard methods of statistics, e.g. statistical tests (for instance: "Student's t-test" or "Hotelling's multivariate T2 test") and thus justify the use thereof.

The methylation signal has the property that the hybridization noise shows a nearly constant variance over the full range of all possible methylation levels. It is however meaningful for a limited degree only, in order to make statements about absolute methylation values and comparisons deducted therefrom between different detection oligo families.

The fourth step however, the actual calibration of the now variance-stabilized data, has as an aim the transformation of the methylation signals into methylation rates.

In this step of the calibration, these methylation signals are calibrated to measurement signals of the control DNA (g log scores). These are generated by use according to the invention of non-methylated control DNA and defined portions of methylated control DNA in separate, several times repeated microarray hybridizations. The methylation signals are determined, which are assigned to the calibration values (e.g. 0%, 10%, 90% and 100%). Particularly preferred is the embodiment, wherein the control DNA measurement signals only, which correspond to the calibration values 0% and 100%, are produced or used. By means of "maximum likelihood

(ML)" estimation, calibrated methylation rates can now be obtained. The calibrated methylation signal is thus the methylation rate and quantifies the absolute share of methylated DNA in the mixture of methylated and non-methylated DNA. Whereas up to now, it could only be determined per microarray hybridization, whether an oligonucleotide probe pair was significantly differently methylated in different samples, it is now possible, to make for instance statements about whether these significant differences of for instance 25% are in the high or low methylation level, i.e. whether the difference results from a 95% methylation in the one case and a 70% methylation in the other case, or however from a 45% methylation compared to a 20% methylation.

According to the invention is therefore provided a method for determining methylation rates of DNA samples by means of microarrays, which contain CG and TG oligomers, which is characterized by that

the arrays are hybridized with two calibration standards, which have defined methylation rates;

based on the obtained hybridization values, a calibration curve is determined by means of a suitable calculation method, and

the actual methylation rates of the investigated DNA samples are determined by means of this prepared calibration curve.

This method is preferred, wherein the two calibration standards represent methylation rates of 0% and 100%, respectively.

Further, a method according to claim 18 is preferred, wherein more than two calibration standards having different methylation rates are used.

A particularly preferred embodiment is a method according to claim 18, which is characterized by that the actual methylation rates are determined in a multi-stage calculation process, comprising the following steps:

a) a normalization of the hybridization values is performed, whereby methylation signals are determined,

b) a transformation of these signals is performed with the aim of variance stabilization,

c) the methylation rates are determined by using the calibration standards and a suitable maximum likelihood algorithm.

This method is particularly preferred, if before the normalization, the hybridization values are corrected for the background noise inherent in the measurement method.

The verifications or standards described further above can be used for all methods of the quantitative methylation analysis: amongst others, for MS-SNuPE, for the hybridization on microarrays, hybridization assays in solution, direct bisulphite sequentiation, for real time PCR (e.g. HeavyMethyl, MSP. comp. for the PMR values: Eads et al., CANCER RESEARCH 61, 3410-3418, April 15, 2001).

A particularly preferred use of the DNA produced by WGA methods and of the method according to the invention is the diagnosis of cancer dis-

eases or other diseases associated with a modification of the methylation state. Thereto belong amongst others, CNS malfunctions, aggression symptoms or conduct disorders; clinical, psychological and social consequences of brain damages; psychotic disorders and personality disorders; dementia and/or associated syndromes; cardiovascular disease, malfunction and damages; malfunction, damages or disease of the gastrointestinal tract; malfunction, damages or disease of the breathing system; injury, inflammation, infection, immunity and/or convalescence; malfunction, damages or disease of the body as a deviation in the development process; malfunction, damages or disease of the skin, of the muscles, of the connective tissue, or the bones; endocrine and metabolic malfunction, damages or disease; headaches or sexual malfunction. The method according to the invention is further suitable for the prognosis of undesired effects of drugs and for the differentiation of cell types or tissues or for the investigation of the cell differentiation.

According to the invention is furthermore provided a kit composed of reagents for performing a WGA method or of DNA amplified already by a WGA method and of reagents for performing a bisulphite conversion, and optionally also contains a polymerase, primers and/or probes for an amplification and detection.

According to the invention is further provided fully methylated DNA, which has been produced by a WGA method and then methylated by means of an enzyme, preferably the SssI methylase. According to the invention is finally also

provided a mixture of methylated and non-methylated DNA produced by a genome-wide amplification method. The use of such a mixture in particular as a standard in the methylation analysis is an essential part of the present invention. It is preferred that the production of the mixture is performed as described above.

Preferred are mixtures with a share of 1%, 2%, 5%, 10%, 25%, 50% methylated DNA, as well as mixtures with a share of 99%, 98%, 95%, 90%, 75%, 50%. For the use in the microarray application, besides the use of the standards 0% and 100%, rather low-concentrated mixtures, i.e. 1%, 2%, 5%, 10%, 25% and 50%, are preferred. Preferred is further the use of mixtures in a high concentration range, i.e. 99%, 98%, 95%, 90%, 75% and 50%.

Particularly preferred is however also the advantageous simultaneous use of mixtures in a high concentration range and in a low concentration range.

Preferred is furthermore the use of 100% non-methylated and 100% methylated DNA.

For the use in the "sensitive detection" range, for instance with assays such as HeavyMethyl and MSP, it is preferred that the following mixture ratios are used: 1:10,000, 1:5,000, 1:2,500, 1:1,000, 1:500, 1:200, 1:100. For MSP assays correspondingly smaller concentrations are preferably used, such as 1:10,000, 1:5,000, 1:2,500, 1:1,000, 1:500, 1:200, 1:100.

Examples:

Example 1: use of MDA-DNA for calibrations.

The methylation degree of a DNA from abdominal fatty tissue is to be determined by means of oligonucleotide microarrays and for a comparison by means of the MS-SNuPE method. For this purpose, the DNA is extracted from the biological sample by means of the QIAamp Mini Kit (Qiagen) according to manufacturer's instructions. For determining a calibration curve, different mixtures of methylated and non-methylated DNA are produced (0%, 25%, 50%, 75%, 100% methylated DNA). The non-methylated DNA was obtained from Molecular Staging, where it has been produced by an MDA reaction from human genomic DNA of full blood. In an MDA reaction, all methylation signals are deleted (see above). The fully methylated DNA is produced from the MDA-DNA by means of an SssI methylase (New England Biolabs). The synthesis is performed according to manufacturer's instructions. The remaining non-methylated DNA is reacted with all reagents except for the SssI methylase. Thus it is secured that the DNA concentration is identical in both batches, and that in both batches the same reaction components are present. Then, non-methylated and methylated DNA are mixed in the following ratios: 4:0 for 0%, 3:1 for 25%, 2:2 for 50%, 1:3 for 75%, 0:4 for 100%. The DNA is then bisulphite-converted in presence of dioxane as a denaturating solvent (cf.: DE 10029 915 Al; German application: File No.: 10347396.3). Subsequently, the prepared DNA mixtures and the DNA from the biological sample are employed in a multiplex PCR. 8 fragments each are amplified.

As primers are used the oligonucleotides listed in Table 1. The amplifications are performed by means of the QIAGEN HotStarTaq Kit essentially according to manufacturer's instructions and with the following temperature profile: 95°C: 15 min; 45 times: (95°C: 15 sec; 55°C: 30 sec; 72°C: 60 sec); 72°C: 10 min. The multiplex PCR products are then hybridized at an oligomer microarray. The probe oligonucleotides are listed in Table 2. The hybridization and the methylation signal determination are made as described by Adorjan et al., 2002 (ibid.). For each sample and each calibration mix, eight hybridizations are performed. For the preparation of calibration curves for a CpG position, the measured methylation rate is plotted against the theoretical methylation rate. The measured methylation rate results from the signal intensity of an oligonucleotide probe, which is specific for the methylated state, divided by the total intensity of this probe + a matching (i.e. covering the same CpG position) probe, which is specific for the non-methylated state. The theoretical methylation state corresponds to the methylation levels of the used defined mixtures. Oligonucleotide probe pairs, which are suitable for calibration purposes, have monotonously increasing calibration curves. For the Ms-SNuPE reaction, the samples are amplified with the primers mentioned above in individual PCR reactions. The reaction conditions are the same as for the multiplex PCR (see above). In the extension reaction, positions are used as primer binding sites, which are positioned directly at the flanks of CpG positions, which correspond to those of the oligonucleotide microarrays. The Ms-SNuPE assay is performed according to the in-

structions of the manufacturer of the MegaBace-SNuPE kit. For the two possible variants of the nucleotides to be incorporated, ddNTP's marked with different dyes are used. For every SNuPE assay, four measurements are made in parallel. The signal intensities determined by the SNP profile software (Amersham) (Imeth- for non-methylated specific probes and Imeth+ for methylated specific probes) of the two employed dyes are used according to the quotient Imeth+/(Imeth- + Imeth+), in order to determine the measured methylation rate. By plotting these values against the theoretical methylation rate, again a calibration curve is obtained, which should be monotonously increasing. The monotonously increasing calibration curves thus generated are used to determine the actual methylation from the measured methylation rate of sample DNA.

The methylation rates in sample DNA determined by the two methods microarray analysis (chip) and SNuPE and corrected at corresponding calibration curves are in a good agreement for the shown CpG positions. These data show that non-methylated DNA produced by MDA or corresponding mixtures with methylated DNA can be used very well as a standard in the methylation analysis.

Example 2:

The results of an analogously performed experiment with however partially differing oligo sequences are shown in Fig. 1. The y-axis repre-

sents the percentage of methylation, the x-axis shows the hybridization at different oligonucleotides or different SNuPE assays.

Figures.

Figure 1:

Fig. 1 shows the results of an experiment described analogously to Example 1. The y-axis shows the determined percentage of methylation, the x-axis indicates that different amplificates (1 - 5) have been investigated. The various hatchings indicate, by which method this value has been determined. The hybridization at different immobilized oligonucleotides is indicated herein as a "chip", and the alternatively used method is the SNuPE assay, designated "SNuPE". Further is indicated, whether it is DNA methylated up to 100% (100%), 0% methylated, i.e. non-methylated DNA (0%) or naturally occurring DNA from the human abdominal tissue. The methylation rates in sample DNA determined by the two methods and corrected at respective calibration curves are in a good agreement for the shown CpG positions.

Figs. 2 and 3 show plots, which indicate at the y-axis the absolute methylation rates (methylation levels) in breast cancer samples and lymphocyte samples. For each detection oligo, minimum and maximum hybridization intensities have been determined. This was achieved by hybridization with fully non-methylated human DNA (Molecular Staging, New Haven, CT) on the one

hand, and enzymatically methylated control DNA (SssI; New England Biomedical), as calibration standards on the other hand. The share of the methylation of a sample at a certain CpG position is determined by linear interpolation between the corresponding minimum intensity (0% methylation of CG oligos, 100% methylation for TG oligos) and the maximum intensity (100% methylation of CG oligos, 0% methylation for TG oligos). The linear interpolation is performed by application of a "maximum likelihood algorithm", which takes into account the hybridization-specific error distribution (for this see Rocke and Durbin, (2001) A model for measurement error for gene expression arrays, Journal of Computational Biology, 8, 557-569). Likelihoods of CG and TG oligomers of the same CpG position are combined in a value for the methylation share. Figs. 2 and 3 show the distribution of determined methylation rates for every CpG position, sorted according to the median of the methylation rate.

In Fig. 2 are shown results of lymphocyte samples.

In Fig. 3 are shown results of breast cancer data.

Table 1: Primers for the multiplex amplification

Gene Name	RefSeq-ID	Primer orientation	Sequence
ERS1	NM_000125 & NM_002920	forward	AGGAGGGGGAATTAAATAGA

		reverse	ACAATAAAACCATCCCAAATAC
AR	NM_000044	forward	GTAGTAGTAGTAGTAAGAGA
		reverse	ACCCCCTAAATAATTATCCT
CDKN2a	NM_000077	forward	GGGGTTGGTTGGTTATTAGA
		reverse	AACCCTCTACCCACCTAAAT
CDKN2B	NM_004936	forward	GGTTGGTTGAAGGAATAGAAAT
		reverse	CCCACTAAACATACCCTTATTC
GSTP1	NM_000852	forward	ATTTGGGAAAGAGGGAAAG
		reverse	TAAAAACTCTAAACCCCATCC
TP73	NM_005427	forward	AGTAAATAGTGGGTGAGTTATGAA
		reverse	GAAAAACCTCTAAAAACTACTCTCC
MLH1	NM_000249	forward	TAAGGGGAGAGGAGGAGTTT
		reverse	ACCAATTCTCAATCATCTCTTT
MGMT	NM_002412	forward	AAGGTTTTAGGGAAGAGTGTTT
		reverse	ACCTTTTCCTATCACAAAAATAA

Table 2: Oligonucleotide probes

Name of olignucleotide	Sequence
Oligonucleotide probes for ERS1	
ERS1:204A209	ATTTAGTAGCGACGATAAGT
ERS1:204A204	GTAGCGACGATAAGTAAAGT
ERS1:204A217	TTAGTAGCGACGATAAGTAAA
ERS1:204A2212	TTTATTTAGTAGCGACGATAAG
ERS1:204B237	TTAGTAGTGATGATAAGTAAAGT
ERS1:204B2413	TTTTATTTAGTAGTGATGATAAGT
ERS1:204B2512	TTTATTTAGTAGTGATGATAAGTAA
ERS1:204B2511	TTATTTAGTAGTGATGATAAGTAAA
ERS1:248A195	GGGATCGTTTTAAATCGAG
ERS1:248A204	GGATCGTTTTAAATCGAGTT

Name of oligonucleotide	Sequence	
ERS1:248A206	TGGGATCGTTTTAAATCGAG	
ERS1:248A213	GATCGTTTTAAATCGAGTTGT	
ERS1:248B216	TGGGATTGTTTTAAATTGAGT	
ERS1:248B223	GATTGTTTTAAATTGAGTTGTG	30
ERS1:248B224	GGATTGTTTTAAATTGAGTTGT	
ERS1:248B228	TTTGGGATTGTTTTAAATTGAG	
ERS1:607A183	GTTCGCGGTTACGGATTA	
ERS1:607A193	GTTCGCGGTTACGGATTAT	
ERS1:607A194	TGTTCGCGGTTACGGATTA	
ERS1:608A203	GTTCGCGGTTACGGATTATG	
ERS1:607B213	GTTTGTGGTTATGGATTATGA	
ERS1:607B219	TATTTTGTTTGTGGTTATGGA	
ERS1:607B215	TTGTTTGTGGTTATGGATTAT	
ERS1:607B227	TTTTGTTTGTGGTTATGGATTA	40
ERS1:451A193	TATCGGATTCGTAGGTTTT	
ERS1:451A204	TTATCGGATTCGTAGGTTTT	
ERS1:451A206	TTTTATCGGATTCGTAGGTT	
ERS1:451A207	GTTTTATCGGATTCGTAGGT	
ERS1:451B218	GGTTTTATTGGATTTGTAGGT	
ERS1:451B226	TTTTATTGGATTTGTAGGTTTT	
ERS1:451B227	GTTTTATTGGATTTGTAGGTTT	
ERS1:451B237	GTTTTATTGGATTTGTAGGTTTT	
Oligonucleotide probes for AR		
AR:971A188	AGTATTTTCGGACGAGGA	
AR:971A183	TTTCGGACGAGGATGATT	50
AR:971A196	TATTTTCGGACGAGGATGA	
AR:971A1910	TTAGTATTTTCGGACGAGG	
AR:971B218	AGTATTTTTGGATGAGGATGA	
AR:971B2112	TGTTAGTATTTTTGGATGAGG	

Name of oligonucleotide	Sequence	
AR:971B213	TTTTGGATGAGGATGATTTAG	
AR:971B217	GTATTTTTGGATGAGGATGAT	
AR:1137°164	GTAGCGGGAGAGCGAG	
AR:1137°175	AGTAGCGGGAGAGCGAG	
AR:1137°186	TAGTAGCGGGAGAGCGAG	
AR:1137B183	TAGTGGGAGAGTGAGGGA	60
AR:1137B185	AGTAGTGGGAGAGTGAGG	
AR:1137B197	GTAGTAGTGGGAGAGTGAG	
AR:1137B174	GTAGTGGGAGAGTGAGG	
AR:869A195	ATAGTCGTAGTCGGTTTTG	
AR:869A208	TTTATAGTCGTAGTCGGTTT	
AR:869A219	TTTTATAGTCGTAGTCGGTTT	
AR:869A2111	ATTTTTATAGTCGTAGTCGGT	
AR:869B193	AGTTGTAGTTGGTTTTGGA	
AR:869B2212	AATTTTTATAGTTGTAGTTGGT	
AR:869B2313	TAATTTTTATAGTTGTAGTTGGT	70
AR:869B2414	GTAATTTTTATAGTTGTAGTTGGT	
AR:814A228	AAGTTTATCGTAGAGGTTTTAT	
AR:814A2212	TTTTAAGTTTATCGTAGAGGTT	
AR:814A2310	TTAAGTTTATCGTAGAGGTTTTA	
AR:814A238	AAGTTTATCGTAGAGGTTTTATA	
AR:814B228	AAGTTTATTGTAGAGGTTTTAT	
AR:814B2210	TTAAGTTTATTGTAGAGGTTTT	
AR:814B2212	TTTTAAGTTTATTGTAGAGGTT	
AR:814B2211	TTTAAGTTTATTGTAGAGGTTT	
Oligonucleotide probes for CDKN2A		
CDKN2A:2147A173	GGGCGTTGTTTAACGTA	80
CDKN2A:2147A183	GGGCGTTGTTTAACGTAT	
CDKN2A:2147B195	GGGGGTGTTGTTTAATGTA	

Name of oligonucleotide	Sequence	
CDKN2A:2147B194	GGGGTGTTGTTTAATGTAT	
CDKN2A:2157A217	TGTTTAACGTATCGAATAGTT	
CDKN2A:2157A227	TGTTTAACGTATCGAATAGTTA	
CDKN2A:2157A228	TTGTTTAACGTATCGAATAGTT	
CDKN2A:2157A238	TTGTTTAACGTATCGAATAGTTA	
CDKN2A:2157B229	GTTGTTTAATGTATTGAATAGT	
CDKN2A:2157B239	GTTGTTTAATGTATTGAATAGTT	
CDKN2A:2157B249	GTTGTTTAATGTATTGAATAGTTA	50
CDKN2A:2183A176	GGAGGTCGATTTAGGTG	
CDKN2A:2183A186	GGAGGTCGATTTAGGTGG	
CDKN2A:2183B186	GGAGGTTGATTTAGGTGG	
CDKN2A:2172A183	TTACGGTCGGAGGTCGAT	
CDKN2A:2172A165	AGTTACGGTCGGAGGT	
CDKN2A:2172A176	TAGTTACGGTCGGAGGT	
CDKN2A:2172A188	AATAGTTACGGTCGGAGG	
CDKN2A:2172B198	AATAGTTATGGTTGGAGGT	
CDKN2A:2172B209	GAATAGTTATGGTTGGAGGT	
CDKN2A:2172B194	GTTATGGTTGGAGGTTGAT	100
CDKN2A:2172B203	TTATGGTTGGAGGTTGATTT	
Oligonucleotide probes for CDKN2B		
CDKN2B:2279A185	GTTTACGGTTAACGGTGG	
CDKN2B:2279A183	TTACGGTTAACGGTGGAT	
CDKN2B:2279A197	AAGTTTACGGTTAACGGTG	
CDKN2B:2279A209	TTAAGTTTACGGTTAACGGT	
CDKN2B:2279B206	AGTTTATGGTTAATGGTGGA	
CDKN2B:2279B216	AGTTTATGGTTAATGGTGGAT	
CDKN2B:2279B223	TTATGGTTAATGGTGGATTATT	
CDKN2B:2279B2210	GTTAAGTTTATGGTTAATGGTG	
CDKN2B:2330A156	GGAATGCGCGAGGAG	110

Name of oligonucleotide	Sequence	
CDKN2B:2330A165	GAATGCGCGAGGAGAA	
CDKN2B:2330A175	GAATGCGCGAGGAGAAT	
CDKN2B:2330A184	AATGCGCGAGGAGAATAA	
CDKN2B:2330B186	GGAATGTGTGAGGAGAAT	
CDKN2B:2330B196	GGAATGTGTGAGGAGAATA	
CDKN2B:2330B205	GAATGTGTGAGGAGAATAAG	
CDKN2B:2329B206	GGAATGTGTGAGGAGAATAA	
CDKN2B:2234A168	AGAGAGTGCGTCGGAG	
CDKN2B:2234A167	GAGAGTGCGTCGGAGT	
CDKN2B:2234A166	AGAGTGCGTCGGAGTA	120
CDKN2B:2234A176	AGAGTGCGTCGGAGTAG	
CDKN2B:2234B178	AGAGAGTGTGTTGGAGT	
CDKN2B:2234B188	AGAGAGTGTGTTGGAGTA	
CDKN2B:2234B187	GAGAGTGTGTTGGAGTAG	
CDKN2B:2234B198	AGAGAGTGTGTTGGAGTAG	
CDKN2B:2268A193	TGTCGTTAAGTTTACGGTT	
CDKN2B:2268A194	GTGTCGTTAAGTTTACGGT	
CDKN2B:2268A205	AGTGTCGTTAAGTTTACGGT	
CDKN2B:2268A203	TGTCGTTAAGTTTACGGTTA	
CDKN2B:2268B215	AGTGTTGTTAAGTTTATGGTT	130
CDKN2B:2268B216	GAGTGTTGTTAAGTTTATGGT	
CDKN2B:2268B224	GTGTTGTTAAGTTTATGGTTAA	
CDKN2B:2268B225	AGTGTTGTTAAGTTTATGGTTA	
Oligonucleotide probes for MGMT		
MGMT:597A188	GGATTATTCGGGTACGTG	
MGMT:597A184	TATTCGGGTACGTGGTAG	
MGMT:597A186	ATTATTCGGGTACGTGGT	
MGMT:597A196	ATTATTCGGGTACGTGGTA	
MGMT:597B193	ATTTGGGTATGTGGTAGGT	

Name of oligonucleotide	Sequence	
MGMT:597B205	TTATTTGGGTATGTGGTAGG	
MGMT:597B204	TATTTGGGTATGTGGTAGGT	140
MGMT:597B2212	TTTAGGATTATTTGGGTATGTG	
MGMT:621A174	TGTACGTTCGCGGATTA	
MGMT:621A183	GTACGTTCGCGGATTATT	
MGMT:621A185	TTGTACGTTCGCGGATTA	
MGMT:621A184	TGTACGTTCGCGGATTAT	
MGMT:621B217	GTTTGTATGTTTGTGGATTAT	
MGMT:621B224	TGTATGTTTGTGGATTATTTTT	
MGMT:621B223	GTATGTTTGTGGATTATTTTTG	
MGMT:621B225	TTGTATGTTTGTGGATTATTTT	
MGMT:394A197	TTTTGGACGGTATCGTTTA	150
MGMT:394A206	TTTGGACGGTATCGTTTATT	
MGMT:394A208	TTTTTGGACGGTATCGTTTA	
MGMT:394A213	GGACGGTATCGTTTATTATAG	
MGMT:394B2111	TAGTTTTTGGATGGTATTGTT	
MGMT:394B229	GTTTTTGGATGGTATTGTTTAT	
MGMT:394B234	TGGATGGTATTGTTTATTATAGG	
MGMT:394B237	TTTTGGATGGTATTGTTTATTAT	
MGMT:530A173	TTTCGAGTAGGATCGGG	
MGMT:530A184	GTTTCGAGTAGGATCGGG	
MGMT:530A183	TTTCGAGTAGGATCGGGA	160
MGMT:530A193	TTTCGAGTAGGATCGGGAT	
MGMT:530B194	GTTTTGAGTAGGATTGGGA	
MGMT:530B193	TTTTGAGTAGGATTGGGAT	
MGMT:530B203	TTTTGAGTAGGATTGGGATT	
MGMT:530B204	GTTTTGAGTAGGATTGGGAT	
Oligonucleotide probes for MLH1		
MLH1:1753A176	GAAGAGCGGATAGCGAT	

Name of oligonucleotide	Sequence
MLH1:1753A185	AAGAGCGGATAGCGATTT
MLH1:1753A184	AGAGCGGATAGCGATTTT
MLH1:1753A193	GAGCGGATAGCGATTTTTA
MLH1:1753B198	AGGAAGAGTGGATAGTGAT
MLH1:1753B2110	ATAGGAAGAGTGGATAGTGAT
MLH1:1753B214	AGAGTGGATAGTGATTTTTAA
MLH1:1753B226	GAAGAGTGGATAGTGATTTTTA
MLH1:2026A186	AAATGTCGTTCGTGGTAG
MLH1:2026A197	AAAATGTCGTTCGTGGTAG
MLH1:2026A1910	GTTAAAATGTCGTTCGTGG
MLH1:2026A209	TTAAAATGTCGTTCGTGGTA
MLH1:2026B195	AATGTTGTTTGTGGTAGGG
MLH1:2026B207	AAAATGTTGTTTGTGGTAGG
MLH1:2026B218	TAAAATGTTGTTTGTGGTAGG
MLH1:2026B2110	GTTAAAATGTTGTTTGTGGTA
MLH1:1770A186	TTTTAACGCGTAAGCGTA
MLH1:1770A194	TTAACGCGTAAGCGTATAT
MLH1:1770A195	TTTAACGCGTAAGCGTATA
MLH1:1770A203	TAACGCGTAAGCGTATATTT
MLH1:1770B239	GATTTTTAATGTGTAAGTGTATA
MLH1:1770B234	TTAATGTGTAAGTGTATATTTTT
MLH1:1770B249	GATTTTTAATGTGTAAGTGTATAT
MLH1:1770B259	GATTTTTAATGTGTAAGTGTATATT
MLH1:1939A173	GAACGTGAGTACGAGGT
MLH1:1939A183	GAACGTGAGTACGAGGTA
MLH1:1939A185	AAGAACGTGAGTACGAGG
MLH1:1939A186	GAAGAACGTGAGTACGAG
MLH1:1939B207	GGAAGAATGTGAGTATGAGG
MLH1:1939B208	AGGAAGAATGTGAGTATGAG

Name of oligonucleotide	Sequence	
MLH1:1939B216	GAAGAATGTGAGTATGAGGTA	
MLH1:1939B213	GAATGTGAGTATGAGGTATTG	
Oligonucleotide probes for TP73		
TP73:750A174	GATTCGTTGCGGTTAGA	
TP73:750A184	GATTCGTTGCGGTTAGAG	
TP73:750A183	ATTCGTTGCGGTTAGAGA	200
TP73:750A185	GGATTCGTTGCGGTTAGA	
TP73:750B205	GGATTTGTTGTGGTTAGAGA	
TP73:750B213	ATTTGTTGTGGTTAGAGAATT	
TP73:750B214	GATTTGTTGTGGTTAGAGAAT	
TP73:750B223	ATTTGTTGTGGTTAGAGAATTT	
TP73:1082A164	GGTGCGCGTAGAGAAT	
TP73:1082A166	TTGGTGCGCGTAGAGA	
TP73:1082A165	TGGTGCGCGTAGAGAA	
TP73:1082A174	GGTGCGCGTAGAGAATA	
TP73:1082B1810	AGGTTTGGTGTGTGTAGA	210
TP73:1082B195	TGGTGTGTGTAGAGAATAA	
TP73:1082B207	TTTGGTGTGTGTAGAGAATA	
TP73:1082B217	TTTGGTGTGTGTAGAGAATAA	
TP73:858A186	GGATATCGGTTCGGAGTT	
TP73:858A189	AGAGGATATCGGTTCGGA	
TP73:858A195	GATATCGGTTCGGAGTTAG	
TP73:858A193	TATCGGTTCGGAGTTAGAT	
TP73:858B2011	GTAGAGGATATTGGTTTGGA	
TP73:858B208	GAGGATATTGGTTTGGAGTT	
TP73:858B224	ATATTGGTTTGGAGTTAGATTA	220
TP73:858B235	GATATTGGTTTGGAGTTAGATTA	
TP73:1135A204	ATATCGAACGGGATTTAGAG	
TP73:1135A2112	TTTTTTAAATATCGAACGGGA	

Name of oligonucleotide	Sequence
TP73:1135A228	TTAAATATCGAACGGGATTTAG
TP73:1135A229	TTTAAATATCGAACGGGATTTA
TP73:1135B224	ATATTGAATGGGATTTAGAGTT
TP73:1135B237	TAAATATTGAATGGGATTTAGAG
TP73:1135B248	TTAAATATTGAATGGGATTTAGAG
TP73:1135B2413	TTTTTTTAAATATTGAATGGGATT
Oligonucleotide probes for GSTP1	
GSTP1:1900A157	GGGAGTTCGCGGGAT
GSTP1:1900A166	GGAGTTCGCGGGATTT
GSTP1:1900A175	GAGTTCGCGGGATTTTT
GSTP1:1900A185	GAGTTCGCGGGATTTTTT
GSTP1:1900B177	GGGAGTTTGTGGGATTT
GSTP1:1900B187	GGGAGTTTGTGGGATTTT
GSTP1:1900B196	GGAGTTTGTGGGATTTTTT
GSTP1:1900B206	GGAGTTTGTGGGATTTTTTA
GSTP1:2007A196	GAGTTTCGTCGTCGTAGTT
GSTP1:2007B198	TGGAGTTTTGTTGTTGTAG
GSTP1:2007B219	TTGGAGTTTTGTTGTTGTAGT
GSTP1:2126A207	GGTTTTTCGTTTATTTCGAG
GSTP1:2126A216	GTTTTTCGTTTATTTCGAGAT
GSTP1:2126A218	AGGTTTTTCGTTTATTTCGAG
GSTP1:2126A226	GTTTTTCGTTTATTTCGAGATT
GSTP1:2126B217	GGTTTTTTGTTTATTTTGAGA
GSTP1:2126B227	GGTTTTTTGTTTATTTTGAGAT
GSTP1:2126B228	AGGTTTTTTGTTTATTTTGAGA
GSTP1:2126B2310	GTAGGTTTTTTGTTTATTTTGAG
GSTP1:2142A153	ATTCGGGACGGGGGT
GSTP1:2142A154	GATTCGGGACGGGGG
GSTP1:2142A155	AGATTCGGGACGGGG

Name of oligonucleotide	Sequence
GSTP1:2142A156	GAGATTCGGGACGGG
GSTP1:2142B174	GATTTGGGATGGGGTT
GSTP1:2142B175	AGATTTGGGATGGGGT
GSTP1:2142B176	GAGATTTGGGATGGGG
GSTP1:2142B184	GATTTGGGATGGGGTTT
Oligonucleotide probes for CDH13 C	
CDH13:137A1810	ATGTTATTTTCGCGGGGT
CDH13:137B1910	ATGTTATTTTTGTGGGGTT
CDH13:137B2011	GATGTTATTTTTGTGGGGTT
CDH13:153A174	TTTTCGCGAGGTGTTTA
CDH13:153A184	TTTTCGCGAGGTGTTTAT
CDH13:153A185	TTTTTCGCGAGGTGTTTA
CDH13:153A195	TTTTTCGCGAGGTGTTTAT
CDH13:153B206	GTTTTTTGTGAGGTGTTTAT
CDH13:153B215	TTTTTTGTGAGGTGTTTATTT
CDH13:153B216	GTTTTTTGTGAGGTGTTTATT
CDH13:153B226	GTTTTTTGTGAGGTGTTTATTT
CDH13:187A173	AAACGAGGGAGCGTTAG
CDH13:187A174	AAAACGAGGGAGCGTTA
CDH13:187A175	TAAAACGAGGGAGCGTT
CDH13:187A185	TAAAACGAGGGAGCGTTA
CDH13:187B183	AAATGAGGGAGTGTTAGG
CDH13:187B193	AAATGAGGGAGTGTTAGGA
CDH13:187B194	AAAATGAGGGAGTGTTAGG
CDH13:187B197	TGTAAAATGAGGGAGTGTT
CDH13:22°203	GGTCGTTAGTTTTTCGTGTA
CDH13:22B203	GGTTGTTAGTTTTTTGTGTA
CDH13:22B213	GGTTGTTAGTTTTTTGTGTAA
CDH13:22B214	TGGTTGTTAGTTTTTTGTGTA

Name of oligonucleotide	Sequence
CDH13:22B223	**GGTTGTTAGTTTTTTGTGTAAT**

780

All documents and references mentioned in this specification are herewith included in their full scope in the scope of disclosure of the present invention.